Trends in Neuroscience and Education 43 (2026) 100290

Contents lists available at ScienceDirect

Trends in Neuroscience and Education

journal homepage: www.elsevier.com/locate/tine

Research paper

The sound of the roman empire: effects of playing a history video game with and without sound

Amedeo Viccari [a,b,*], Marlit Annalena Lindner [c,d], Anna Kienitz [a], Jens-Uwe Hahn [e], Richard Göllner [f], Valentin Emslander [a], Lisa Bardach [a,g]

[a] Department of Psychology, University of Giessen, Giessen, Germany
[b] Hector Research Institute of Education Sciences and Psychology, University of Tübingen, Tübingen, Germany
[c] Europa-Universität Flensburg, Flensburg, Germany
[d] Leibniz Institute for Science and Mathematics Education (IPN), Kiel, Germany
[e] Stuttgart Media University (HdM), Stuttgart, Germany
[f] University of Potsdam, Potsdam, Germany
[g] Kaunas University of Technology, Kaunas, Lithuania

ARTICLE INFO

Keywords:
Game-based learning
History learning
Academic outcomes
Motivation
Academic emotions
Video games

ABSTRACT

Background: Sound is a core component of digital games, and its integration is assumed to support learning, motivation, and positive emotions. However, empirical evidence on the role of sound in educational video games remains limited, particularly in narrative-driven educational adventure games such as digital history games.

Methods: In a laboratory experiment, university students (N = 111) either played an educational history video game without sound or with additional sound features (ambient audio, character voices, and narrated codex entries providing additional historical information). Post-test measures assessed factual knowledge, triggered and maintained situational interest, and academic emotions (enjoyment and boredom). Engagement with optional supplemental historical information provided through in-game codex entries was measured using behavioral log data. We analyzed differences between the two conditions while controlling for relevant pre-test variables.

Results: Participants in the add-on sound condition did not score significantly higher on the knowledge test than those in the no-sound condition. Likewise, no statistically significant differences emerged in situational interest, enjoyment, boredom, or codex engagement between conditions. Additional analyses indicated that participants' interactions with codex entries positively predicted knowledge test performance, indicating that voluntary engagement with supplemental content contributed to learning.

Conclusion: Our findings suggest that the presence of sound alone may not enhance academic outcomes in a narrative-driven educational video game. Additionally, our findings indicate that learning outcomes depended strongly on learners' engagement with in-game codex entries. Overall, our results on the inclusion of sound highlight the importance of examining specific design features within educational history video games.

Sound is a central component of digital environments and has been assumed to play an important role in shaping learning experiences and outcomes. In multimedia learning research, the inclusion of audio elements is grounded in the Cognitive Theory of Multimedia Learning (CTML; [1]), which proposes that people process information through two limited-capacity channels, one visual-pictorial and one auditory-verbal. When instructional materials distribute information across these channels, learning can be supported by reducing overload on any single channel and facilitating more efficient working-memory use. Interactive computer-based learning environments, including video games, are well-suited to leverage multimedia learning principles, as they often combine visual and auditory information within interactive tasks [2–4]. In line with this, a key principle derived from CTML is the modality principle, which shows that narration can enhance understanding and retention compared to on-screen text, particularly when visual information is complex [5,6]. According to Mayer's multimedia learning framework, coordinated narration and visual information should improve retention and transfer, highlighting how

* Corresponding author at: University of Giessen, Department of Psychology, Otto-Behaghel-Strasse 10F 35390 Giessen.
E-mail address: amedeo.viccari@uni-giessen.de (A. Viccari).

https://doi.org/10.1016/j.tine.2026.100290
Received 13 February 2026; Received in revised form 17 May 2026; Accepted 20 May 2026
Available online 21 May 2026
2211-9493/

audio can support coherent mental model formation [1], which may, in turn, facilitate *knowledge acquisition* in video game-based learning [7].

In addition to cognitive processing, sound has been proposed to influence *affective* and *motivational aspects of learning*. The present study focuses on academic emotions (i.e., enjoyment, boredom [8,9]) and situational interest [10]. However, since these outcomes have rarely been examined in research on video games and sound effects, we next review the broader literature on affective and motivational outcomes. Audio elements—such as background music, environmental sound, and character voices—are integral design features that may affect emotional responses, sense of immersion, and engagement. For instance, musical scores can affect selected perceptual evaluations of the video gaming experience (e.g., how colorful, dangerous, relaxed, or simple the game is perceived to be) [11]. The presence of game sound has also been shown to affect subjective gameplay experience in a first-person shooter game, including immersion, flow, positive affect, and perceived competence [12]. In game-based learning environments, non-player characters (NPC) voiceovers have also been shown to increase learners' self-reported engagement [13]. At the same time, research incorporating music as a design element in educational games indicates that the impact of game audio may depend on how it is implemented and integrated, rather than on its mere presence. For example, Kao et al [14] found that, in an educational programming game, a self-similar avatar voice, compared with a dissimilar avatar voice, increased participants' perceived competence, immersion, time spent playing, and in-game performance, but did not significantly affect enjoyment. Although this manipulation involved personalized voice output rather than ambient sound design, it suggests that audio features embedded in educational games may influence emotional and motivational outcomes.

Sound also plays a role in directing attention, focus, and information selection, and shapes actual behaviors in interactive environments [6, 13]. Specifically, the signaling principle as deduced from CTML proposes that audio cues can help learners direct their attention to important information [5]. Conceptual work on audio game design likewise emphasizes that sound can structure players' spatial orientation, guide navigation, and support attentional focus within interactive environments [15]. From this perspective, audio does not merely accompany visual information but functions as an informational layer that scaffolds perception, facilitates navigation, and shapes how players interact with the game world. Thus, the inclusion of sound may also influence learners' *in-game behaviors*, such as engagement with optional or supplemental content, by enhancing immersion and supporting focused engagement.

Despite strong theoretical arguments and some promising empirical insights, research on the role of sound in educational video games remains limited. Whereas many studies focus on visual design, narrative structure, or gameplay mechanics, comparatively few have examined whether audio elements affect knowledge outcomes, motivation, emotional experiences, or in-game behaviors. This gap may be particularly relevant for subjects such as history: For history learning, immersion, emotional engagement, and the construction of coherent mental models are central to how learners make sense of historical narratives and environments [16]. Narrative-oriented analyses further suggest that sound is a key factor in establishing coherent fictional story worlds and fostering immersion in narrative game environments [17]. Such immersive narrative experiences are often considered relevant to learner engagement in educational history games [18]. Accordingly, empirical research systematically investigating the role of sound in educational history games is needed to gain deeper insights into audio design elements that may facilitate learners' knowledge acquisition, interest, emotional responses, and engagement with historical content.

1. The present study

The present study examined the effects of sound on knowledge test performance, motivational outcomes (triggered and maintained situational interest), academic emotions (enjoyment and boredom), and in-game behavior (engagement with codex entries providing historical information) in *Limes* [19], an educational video game on the Roman Empire. In a between-subjects design, participants were randomly assigned to one of two conditions: a version of the game with add-on sound or a version without sound. All learners engaged with the game for up to 90 min and subsequently completed the post-test measures. All hypotheses, measures, covariates, exclusion rules, and statistical analyses were preregistered prior to conducting the analyses (blinded link for peer-review: https://aspredicted.org/xa6gt7.pdf).

Based on the theoretical frameworks and empirical findings reviewed above, we specified four main hypotheses. First, we hypothesized that participants in the add-on sound condition would achieve higher scores on the post-test knowledge measure than those in the no-sound condition, as the inclusion of audio narration and ambient sound may support cognitive processing and the formation of coherent mental models [1] (*Hypothesis 1, H1*).

Second, we expected participants exposed to sound to report higher levels of situational interest (*Hypothesis 2, H2*). Although prior research on sound effects in educational video games that directly measures situational interest is, to the best of our knowledge, lacking, there is some evidence on effects of background music and narrative elements on learners' motivation [11,14]. Drawing on the interest development framework proposed by Hidi and Renninger [10], sound effects are expected to *trigger* situational interest by increasing the perceptual salience and experiential vividness of the learning environment. Insofar as sound effects help sustain engagement with task-relevant information across the learning episode, they may also support *maintained* situational interest.

Third, building on prior research showing positive effects on learners' emotional experiences [11,12], we hypothesized that learners in the add-on sound condition would report more adaptive academic emotions—specifically, higher enjoyment and lower boredom—compared with those in the no-sound condition (*Hypothesis 3, H3*).

Finally, we expected that participants in the add-on sound condition would more often explore additional historical information (referred to in the game as codex entries) and for a longer duration (*Hypothesis 4, H4*), reflecting the role of sound in guiding attention and information selection within interactive learning environments [6,13].

2. Methods

2.1. Participants and design

A total of 112 university students enrolled in psychology, teacher education, and sport science programs participated in the study (84 women, 28 men; *M*age= 23.19, *SD* = 3.95). Data collection was carried out in a controlled laboratory setting at a German university. Based on an a priori power analysis using G*Power 3.1 [20] for an analysis of covariance comparing two independent groups, we aimed to recruit at least $N = 102$ participants, assuming a medium effect size ($f = 0.25$), an alpha level of .05, and a power of 0.80. To account for potential exclusions, we collected data from 112 individuals. The study received ethics approval from the University of [blinded for peer-review] ethics board (approval code: [blinded]). Fig. 3 presents an overview of the study procedure.

Participants were recruited through multiple channels, including campus flyers, university mailing lists, and the university's participation platform, where the study was listed as a standard in-lab study offering course credit as compensation. To be eligible, participants had to be at least 18 years old, fluent in German, not enrolled in a history degree program, and not have uncorrected severe visual impairments.

Participants were randomly assigned to play the video game either with or without sound. We collected data from 76 participants and included 36 participants from a previous study for the no-sound condition. The additional participants were recruited within the same data collection period, in the same laboratory setting, and following an identical procedure, including recruitment, instructions, and technical setup. The previous study also employed random assignment to different conditions and focused on a different research question (a comparison of the *Limes* game and a parallel text format, and of two instructions: having fun versus learning), but did not involve sound [21]. We included only participants from the game condition of the previous study who had received the learning instruction to align with the participant instruction in the current study. Overall, 56 participants (40 women, 16 men) were in the *add-on sound* condition, and 56 participants (44 women, 12 men) were in the *no-sound* condition.

2.2. Experimental manipulation and materials

In our between-subjects design (sound versus no sound), participants engaged with *Limes*, an educational 2D top-down adventure game developed in Unity (version 2021.3.7f1) [19].

In the add-on sound condition, ambient sounds, character voices, and narrated codex entries were enabled. Ambient sounds consisted of continuous environmental audio (e.g., water, wind, and background environmental effects) that accompanied exploration but were not directly linked to instructional content. Character voices and codex narrations presented spoken versions of the same text that was simultaneously displayed on screen. The on-screen text was fully visible from the beginning of each dialogue or codex entry and was not progressively revealed in synchrony with the narration. Thus, auditory and visual-verbal information were presented concurrently, with narration functioning as an additional layer rather than replacing on-screen text.

In the no-sound condition, these audio elements were disabled, while all other game features remained identical. This manipulation allowed us to isolate the role of audio cues on engagement and learning. Prior to engaging with Limes, all participants received the instruction that they should use the video game to learn as much as possible about the history content.

Limes presents a historically grounded narrative set along the Roman frontier during the third century CE. Players follow the journey of a Germanic child who is taken by Roman soldiers and gradually forms an unexpected friendship with a Roman centurion, Ariovist. The story unfolds across five levels, each containing a main storyline, optional secondary missions, and supplementary codex entries that provide additional historical context (see Fig. 1). Codex entries (see Fig. 2 for an example) are short and contain texts covering topics such as Roman emperors, military structures, and everyday life along the limes (the Roman Empire's border).

The game includes 36 codex entries distributed across the five levels. Players can voluntarily open these entries by interacting with characters or objects in the game world. Each entry consists of a short text; in the German version used in this study, entries averaged 148 words (range 59–265), totaling 6328 words of codex material, and were unlocked through interactions with in-game characters or objects. Across all participants, behaviors were automatically recorded, including time spent reading codex entries, exploration of secondary missions, and overall time-on-task (see Table 1 for descriptive engagement measures).

Before beginning the task, participants viewed a short instruction screen that explained the study procedure and encouraged them to engage thoroughly with the available content. This screen remained visible until the experiment leader manually launched the game, and the time spent on the instruction screen was recorded.

All content was designed and reviewed with the aim of providing accurate historical material in an interactive format appropriate for university students. A detailed description of the narrative design, structural elements, and development process of *Limes* can be found in Viccari et al [19].

2.3. Measures

Pre-Test Background Variables and Covariates. In the pre-test, we assessed several background variables as potential covariates for the analyses. Prior knowledge of the learning content was assessed using a short 5-item version of the knowledge test administered at pre-test (possible range: 0–5). Prior interest in history was measured using three items adapted from Schumacher et al [22] (e.g., "How interested are you in Roman history?", rated on a 7-point Likert scale (α = .73). Prior self-reported historical knowledge ("How much knowledge do you have about the following topics?"), was measured with three items, each rated on a 7-point Likert scale (1 = no knowledge at all, 7 = a great deal of knowledge). The items assessed self-reported knowledge of Roman,

Fig. 1. Illustration of codex entry access (top) and secondary mission activation (bottom) in the Limes video game.
Note. The top panel shows how players open a codex entry after interacting with a character. The bottom panel shows the transition from dialogue into a secondary mission, where the player enters an interactive gameplay sequence.

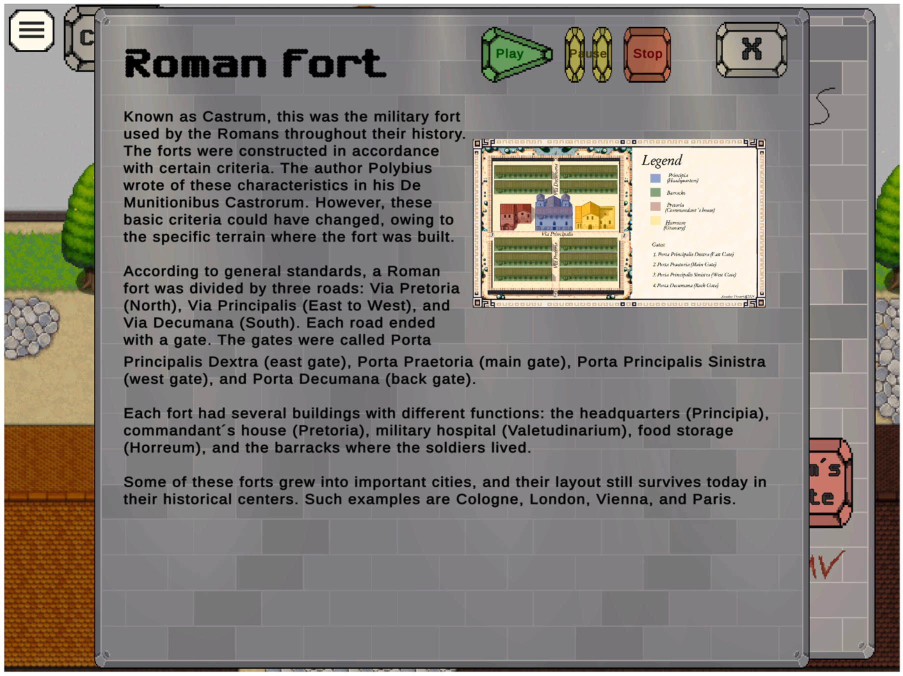

Fig. 2. Codex entry as shown in the video game.

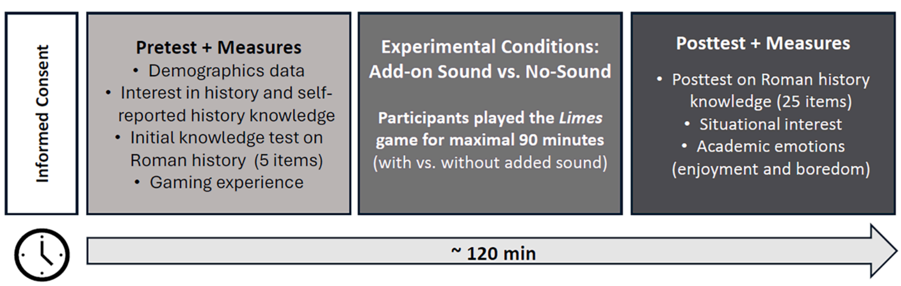

Fig. 3. Overview of the Experimental Procedure.
Note. Maximum completion time maximum was 120 min. In-game behavioral data (time in the codex, codex clicks) was collected while participants played *Limes*.

Table 1
Time spent in limes game (in Minutes) by condition.

	No-Sound (n = 56)		Add-on Sound (n = 55)	
	M	*SD*	*M*	*SD*
Instruction Screen Time (Before Game)	1.62	0.46	1.64	0.29
Codex Entry Time in Game	9.82	9.54	10.15	8.16
Game Play Time	64.03	12.96	65.85	11.66
Total Time Spent in Game	73.85	13.73	76.00	11.56
Post-Test Total Time	9.35	2.29	9.32	1.62

ancient, and general history (adapted from Schumacher et al [22]). Responses were averaged to form a composite scale score (α = .78). Participants also reported their average weekly gaming hours.

Knowledge Test. The knowledge test assessed factual knowledge of Roman history, including topics such as Roman society, military structures, frontier life along the Roman border (limes), and key historical events and figures covered in the learning materials. It consisted of 25 single-choice questions (with 4 answer options, plus an "I do not know" option) that were coded as correct or incorrect. A total sum score was calculated based on the number of correctly answered items (possible range: 0–25). Internal consistency was high (α_{sound} = .79, $\alpha_{no\text{-}sound}$ = .81).

Situational Interest. Situational interest was measured using eleven items adapted from Linnenbrink-Garcia et al [23]. The scale included three components: triggered situational interest (4 items; e.g., "Playing *Limes* is exciting"), maintained interest feeling (4 items; e.g., "I am enthusiastic about what I learn in *Limes*"), and maintained interest value (3 items; e.g., "What I learn in *Limes* is valuable to me"). All items were rated on a 7-point scale (1 = does not apply at all, 7 = fully applies). Following Krebs et al [24], the two maintained interest components were combined into a single maintained-interest scale for further analysis. Internal consistency was high for both triggered interest (α = .86 overall; α_{sound} = .85; $\alpha_{no\text{-}sound}$ = .88) and maintained interest (α = .89 across all participants; α_{sound} = .87; $\alpha_{no\text{-}sound}$ = .91).

Academic Emotions. Academic emotions were assessed using the short version of the Achievement Emotions Questionnaire (AEQ-S; [25]). The measure consists of four items assessing enjoyment (e.g., "Engaging with the content of *Limes* brings me joy") and four items assessing boredom (e.g., "The content of *Limes* is so boring that I catch myself daydreaming"). Participants responded on a 5-point Likert scale (1 = do not agree at all, 5 = fully agree). Internal consistency was acceptable to high for enjoyment (α = .80 overall, α_{sound} = .73; $\alpha_{no\text{-}sound}$ = .86), and boredom (α = .82 overall, α_{sound} = .73; $\alpha_{no\text{-}sound}$ = .87).

All measures used in our study are reported in the Online Supplement.

2.4. Procedure

The experiment took place in a laboratory setting at a German university. Upon arrival, participants were welcomed by the experiment leader, informed about the study procedures and data protection, and they provided informed consent. Participants then completed the pre-test on Unipark via Firefox on university-provided computers. The pre-test included demographic information, self-reported history knowledge, self-reported interest in history, and a short knowledge test on Roman history. After the pre-test, participants received a short instruction screen explaining the task and encouraging them to engage thoroughly with the game content. Assignment to the audio condition (sound vs. no-sound) was determined by randomization embedded in the desktop applications: in the add-on sound condition, ambient audio, character voices, and voiceover of the codex entry were enabled; in the no-sound condition, these audio elements were disabled while all other features remained identical. Once participants had read the instruction screen, the experiment leader manually launched the game application for them. Participants were instructed to spend a maximum of 90 min engaging with the game and wore headphones during the session. Time spent on the instruction screen and all in-game behaviors (e.g., codex engagement, secondary missions, total playtime) were automatically recorded. After the learning phase—or once the allocated time had expired—participants returned to the browser to complete the post-test on Unipark, which included measures of situational interest, academic emotions (enjoyment and boredom), and the 25-item factual knowledge test on Roman history. Each laboratory session lasted up to 120 min in total.

2.5. Analyses

We conducted one-way analyses of covariance (ANCOVAs) to examine the effect of the audio condition (sound vs. no-sound) on all outcome variables. Correlations between pre-test variables and dependent variables were inspected to determine their inclusion as covariates. All significance tests were evaluated at the .05 level, and all directional hypotheses were assessed using one-tailed tests.

3. Results

3.1. Preliminary analyses

As preregistered, participants were to be excluded if their total time spent engaging with Limes fell below 20 min. No participant met this exclusion criterion. One participant was excluded due to a procedural deviation that prevented correct data recording, resulting in a final sample of $N = 111$ for all analyses. On average, participants spent 1.63 min on the instruction screen, followed by approximately 75 min interacting with the game material, and an additional 9.3 min completing the post-test (see Table 1).

To examine whether the two sound conditions differed on pre-test variables, we conducted *t*-tests, which showed no significant differences between the sound and no-sound groups in age $t(108.10) = -0.78$, $p = .44$, weekly gaming hours, $t(69.18) = 0.22$, $p = .83$, prior historical knowledge, $t(108.41) = -0.14$, $p = .89$, interest in history, $t(107.21) = -0.41$, $p = .68$, or pretest factual knowledge, $t(108.95) = -0.08$, $p = .93$. Chi^2 tests likewise indicated no difference in gender distribution across conditions ($\chi^2(1) = 0.51$, $p = .477$) or gaming experience, ($\chi^2(1) = 0.54$, $p = .463$). These results suggest that the two experimental groups were comparable on the pre-test variables assessed. Means and standard deviations for the outcome variables are reported by condition (see Table 2).

To determine whether pre-test characteristics should be controlled for in the main analyses, we inspected the correlations between all pre-test measures and the dependent variables (see Table 3). In this screening, several pre-test variables showed meaningful associations with the outcomes. Accordingly, interest in history, self-reported prior knowledge, and pre-test factual knowledge were included as covariates only in those ANCOVA models in which they showed a significant association with the respective dependent variable.

Table 2
Means and standard deviations for outcome variables.

Outcomes Variables	No-Sound (n = 56)		Add-on Sound (n = 55)	
	M	*SD*	*M*	*SD*
Post-Test Knowledge (Score Range: 0–25)	13.95	4.89	13.80	4.59
Triggered Situational Interest (Scale: 1–7)	4.56	1.26	4.69	1.25
Maintained Situational Interest (Scale: 1–7)	4.28	1.19	4.49	1.04
Enjoyment (Scale: 1–5)	3.39	0.86	3.47	0.69
Boredom (Scale: 1–5)	1.73	0.78	1.63	0.59
Codex Clicks (Count)	15.79	13.34	18.53	11.86
Time in Codex (Minutes)	9.82	9.54	10.15	8.16

3.2. Hypothesis testing

Effects on Situational Interest (Hypothesis 1). The ANCOVA for *triggered situational interest* yielded no significant effect of sound, $F(1108) = 0.22$, $p = .639$, $\eta^2_p < 0.01$. Participants in the add-on sound condition reported triggered interest similar to those in the no-sound condition (Sound: $M = 4.69$, $SD = 1.25$; No-Sound: $M = 4.56$, $SD = 1.26$). Pre-test interest in history was a significant covariate, $F(1108) = 6.42$, $p = .013$, $\eta^2_p = 0.06$. For *maintained situational interest*, the ANCOVA again revealed no significant effect of sound, $F(1107) = 0.85$, $p = .360$, $\eta^2_p = 0.01$. Means were similar between the add-on sound condition ($M = 4.49$, $SD = 1.04$) and the no-sound condition ($M = 4.28$, $SD = 1.19$). Pre-test interest in history significantly predicted maintained interest, $F(1107) = 20.65$, $p < .001$, $\eta^2_p = 0.26$.

Effects on Knowledge Acquisition (Hypothesis 2). The ANCOVA on the 25-item *knowledge post-test* revealed no significant effect of sound on students' knowledge acquisition during the game, $F(1106) = 0.10$, $p = .752$, $\eta^2_p < 0.01$. Participants in the add-on sound condition ($M = 13.78$, $SD = 4.59$) and the no-sound condition ($M = 13.86$, $SD = 4.89$) achieved comparable scores. Pre-test knowledge, $F(1106) = 8.78$, $p = .004$, $\eta^2_p = 0.16$, and pre-test interest in history, $F(1106) = 4.72$, $p = .032$, $\eta^2_p = 0.06$, significantly predicted post-test knowledge scores.

Effects on Academic Emotions (Hypothesis 3). The ANCOVA on *enjoyment* showed no significant effect of sound, $F(1107) = 0.15$, $p = .698$, $\eta^2_p < 0.01$. Enjoyment ratings were comparable across conditions (sound: $M = 3.47$, $SD = 0.69$; no-sound: $M = 3.39$, $SD = 0.86$). Pretest interest in history emerged once again as a significant covariate, $F(1107) = 7.52$, $p = .007$, $\eta^2_p = 0.11$. Finally, the ANCOVA for *boredom* indicated no significant effect of sound, $F(1108) = 0.43$, $p = .513$, $\eta^2_p < 0.01$. Participants in the add-on sound condition reported similar boredom levels ($M = 1.63$, $SD = 0.59$) as those in the no-sound condition ($M = 1.73$, $SD = 0.78$).

Effects on Codex Engagement (Hypothesis 4). The ANCOVA on the number of codex clicks revealed no significant effect of sound, $F(1108) = 1.28$, $p = .260$, $\eta^2_p = 0.01$. Participants in the add-on sound condition opened codex entries at similar rates ($M = 18.51$, $SD = 11.91$) as those in the no-sound condition ($M = 15.77$, $SD = 13.26$). Self-reported prior historical knowledge did not significantly predict codex access, $F(1108) = 0.25$, $p = .616$. A second ANCOVA examined codex engagement time. Again, there was no significant effect of sound, $F(1108) = 0.04$, $p = .844$, $\eta^2_p < 0.01$. Learners in the add-on sound condition spent comparable time in the codex ($M = 10.24$ min, $SD = 8.16$) as those in the no-sound condition ($M = 9.82$ min, $SD = 9.54$). Self-reported prior historical knowledge did not significantly predict codex engagement time, $F(1108) = 0.01$, $p = .927$.

Table 3
Overview of Correlations Between Study Variables.

	1	2	3	4	5	6	7	8	9	10	11
1. Interest in History (Pre-test)	1.00										
2. Self-Reported Knowledge (Pre-test)	.65**	1.00									
3. Pre-Test Knowledge (5 Items)	.41***	.54***	1.00								
4. Weekly Gaming Hours (Pre-test)	.13	.14	.03	1.00							
5. Post-Test Knowledge Test (25 Items)	.33**	.31**	.35**	.09	1.00						
6. Triggered Situational Interest	.27*	.14	−.02	.08	.28*	1.00					
7. Maintained Situational Interest	.52***	.36**	.22	.04	.51***	.73***	1.00				
8. Enjoyment	.42***	.31**	.13	.04	.36**	.84***	.83***	1.00			
9. Boredom	−.20	−.15	−.09	.03	−.30**	−.68***	−.44***	−.56***	1.00		
10. Codex Clicks	.08	.07	.06	−.09	.64***	.24*	.39***	.33**	−.25*	1.00	
11. Time in Codex (Seconds)	.06	.02	.04	−.14	.61***	.33**	.42***	.35**	−.25*	.85***	1.00

Note. *Correlation is significant at the .05 level (two-tailed); **Correlation is significant at the .01 level (two-tailed); ***Correlation is significant at the .001 level (two-tailed).

3.3. Additional analyses (not preregistered)

Bayesian follow-up analyses. To further evaluate the strength of evidence for the null versus the alternative hypothesis (see Hypotheses 1–4), we conducted Bayesian independent-samples *t*-tests for all main outcome variables. The analyses provided moderate evidence in favor of the null hypothesis for time in the codex ($BF_{01} = 3.63$), triggered situational interest ($BF_{01} = 3.25$), enjoyment ($BF_{01} = 3.32$), and boredom ($BF_{01} = 2.97$). Evidence for maintained situational interest ($BF_{01} = 2.43$) and codex clicks ($BF_{01} = 2.15$) was weak and favored the null hypothesis. For post-test knowledge scores, the Bayes factor was inconclusive ($BF_{01} = 0.79$), meaning the evidence was insufficient to support either the null or the alternative hypothesis. Accordingly, conclusions about the effect of sound on knowledge test performance remain tentative. Overall, the Bayesian results converge with the null effects observed in the primary analyses, suggesting that the presence of sound did not have a meaningful impact on learners' interest, emotions, and knowledge test scores.

Effects of Codex Engagement. Following the observation of strong correlations between codex engagement measures and post-test knowledge scores (see Appendix A), we conducted two separate multiple regression analyses predicting knowledge test performance from codex clicks and codex viewing time, respectively, while controlling for relevant covariates. In the first model, codex clicks significantly predicted knowledge scores ($b = 0.21$, $p < .001$), and in the second model, time in the codex significantly predicted knowledge scores ($b = 0.30$, $p < .001$). Prior knowledge was also a significant predictor in both models ($ps < .001$).

4. Discussion

The present study examined how sound in an educational history video game affects knowledge test performance, motivation, academic emotions, and beneficial in-game behaviors (engagement with the codex entries). We compared two versions of the same history video game that differed only in the add-on sound condition, including ambient audio, character voices, and codex narration. Below, we discuss the main findings for each hypothesis, as well as the exploratory findings.

Sound was expected to support learning by distributing processing across auditory and visual channels, thereby reducing visual overload [1]. Contrary to our preregistered hypothesis, however, the inclusion of sound did not lead to higher factual knowledge as assessed by the post-test. Replications of our work are needed before definite conclusions can be drawn; however, we propose that one potential explanation is that *Limes* presents a largely static, text-based informational structure, in which the cognitive load imposed by visual materials may not have been sufficiently high for audio narration to offer an advantage. Prior CTML research shows that modality benefits are most pronounced when learners process fast-paced animations or visually dense demonstrations [5,6], whereas narration offers limited benefits when learners have ample time to process written information. The structure of *Limes*, particularly its self-paced reading segments and extensive textual codex entries, may therefore not have placed learners in conditions where audio would meaningfully lighten cognitive demands.

In addition, the way audio was implemented in *Limes* may have introduced elements of redundancy. Narrated content, particularly in codex entries and dialogues, was presented simultaneously with identical on-screen text. According to the redundancy principle in multimedia learning, such duplication can increase cognitive load because learners must process the same information in both text and audio at the same time [5,26]. In the present study, however, the largely static and self-paced structure of the materials may have limited potential interference between modalities. As a result, narration may neither have provided a clear advantage nor imposed substantial costs, which could explain the absence of differences in knowledge outcomes [27,28].

Beyond cognitive outcomes, sound also did not significantly affect situational interest or academic emotions, in contrast to findings indicating that music and sound effects can enhance immersion, affect, and behavioral engagement [11–13]. As a first potential explanation, the overall very low levels of boredom observed in our study suggest that the game was already perceived as sufficiently engaging, even without audio. The reward of engaging with the game may therefore not have been low in the no-sound condition, leaving limited scope for additional increases in positive activation through sound [29,30]. Another potential explanation is that the audio elements in *Limes*—ambient sounds and sparse voiceovers—may have been relatively subtle compared to continuous, event-linked audio designs often used in studies that reported motivational benefits [12,31]. A further explanation comes from research showing that not all types of audio exert the same motivational influence in educational games [14,32]. For example, Kao et al [14] found that audio can have meaningful motivational effects when it enhances self-relevance and identification with the avatar rather than merely providing background stimulation. Thus, the audio used in future experiments on the effects of voice in history video games may benefit from personalized, identity-linked implementations.

Exploration of codex entries did not differ between the add-on sound and no-sound conditions. This may indicate that codex use in *Limes* is not driven by peripheral sensory features such as ambient audio, but could instead reflect learners' intrinsic motivation and perceived relevance of the task in the game. Research on curiosity and interest supports this pattern: exploratory behaviors are typically initiated when learners perceive informational value or anticipate epistemic rewards, rather than through sensory features alone [33]. Similarly, models of situational interest suggest that whereas environmental features may trigger initial interest, sustained engagement depends primarily on learners' enjoyment of and perceived value in the content [23]. In this context, codex entries provide substantive historical information, and learners may engage with them when they perceive this information as useful for understanding the game world or succeeding in tasks. Interestingly, although we did not assess learners' motivation during gameplay, the bivariate correlations between codex measures and our motivational and emotional outcomes are all positive and moderate in size (see Table 3).

Moreover, additional analyses indicated that codex engagement positively predicted knowledge test performance, even after controlling for prior knowledge. These findings reinforce the instructional value of the codex system and highlight that voluntary engagement with supplemental content appears to be a central mechanism driving learning in Limes, regardless of whether sound is present.

From a design perspective, the present findings suggest that adding audio elements such as narration and ambient sound does not necessarily enhance learning outcomes, particularly when information is already presented in a clear, self-paced textual format. Designers should therefore not assume that additional audio will automatically support cognitive processing, motivation, and emotional responses. Instead, audio may be more effective when it replaces rather than duplicates on-screen text or serves a clear instructional function (e.g., guiding attention or supporting accessibility). At the same time, sound may still contribute to user experience, immersion, and accessibility. Overall, the effectiveness of audio likely depends on how it is functionally integrated into the learning environment rather than on its mere presence.

4.1. Limitations and directions for future research

Several limitations should be acknowledged. First, the sound manipulation may have been too subtle to produce measurable differences. Ambient audio and character voices offer a different experiential profile than continuous narration or directive audio cues, which may be more likely to influence cognitive outcomes or motivational states. Second, the study manipulated the presence versus absence of sound but did not vary the type, intensity, or instructional function of audio. Future research could compare ambient sound, music, fully narrated content, and signaling cues to determine which audio forms offer the strongest learning benefits. Third, the laboratory setting and the use of a single instructional approach may have limited learners' natural engagement with audio. Research on instructional goals suggests that the framing of educational games can influence learners' processing of game content and learning outcomes [34]. Studies using different task instructions may therefore reveal effects that were not observable under the single instructional framing used in the present study.

Additionally, while codex engagement predicted knowledge test performance, the codex entries consisted primarily of written text. Although audio playback was available for every entry, the current study did not track whether participants used this feature. As a result, we cannot determine whether codex audio contributed to learning or whether it was simply underutilized. Future work could develop narrated or partially narrated codex entries with a clearer instructional purpose and examine how different audio formats influence learners who struggle with dense textual material. Adaptive audio support, guided by principles of multimedia learning, may also offer opportunities to enhance comprehension and engagement [6].

5. Conclusions

The present study provides a systematic examination of the role of sound in *Limes*, an educational history video game. The presence of sound did not significantly improve learners' post-test knowledge performance, situational interest, academic emotions and use of supplemental content, as measured by codex engagement. These findings contribute to our understanding of when and whether sound matters in game-based learning.

Financial disclosure

This research was conducted within the LEAD Graduate School & Research Network, funded by the Ministry of Science, Research and the Arts of the state of Baden-Württemberg within the Excellence Initiative II. The work was additionally supported by a Fellowship from the Elite Program for Postdocs of the Baden-Württemberg Foundation and a Jacobs Foundation Research Fellowship (awarded to Lisa Bardach). Lisa Bardach is supported by research funding No P-EDU-23-28, co-funded by the European Union under the 1st April 2025 Agreement with the Research Council of Lithuania.

Ethical statement

The authors confirm that this manuscript has not been published previously and is not under consideration for publication elsewhere. All authors have approved the submission and agree to its publication if accepted. The work complies with Elsevier's Publishing Ethics Policy. The study was conducted in accordance with institutional ethical guidelines and received approval from the responsible ethics committee. The authors declare no conflicts of interest.

CRediT authorship contribution statement

Amedeo Viccari: Writing – original draft, Visualization, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Marlit Annalena Lindner:** Writing – review & editing, Methodology. **Anna Kienitz:** Writing – review & editing, Methodology. **Jens-Uwe Hahn:** Writing – review & editing, Methodology. **Richard Göllner:** Writing – review & editing, Methodology. **Valentin Emslander:** Writing – review & editing, Methodology. **Lisa Bardach:** Writing – review & editing, Supervision, Methodology, Funding acquisition.

Declaration of competing interest

All authors declare that they have no known competing financial interests or personal relationships with other people or organizations that could have inappropriately influenced or biased the work reported in this paper.

Acknowledgments

Amedeo Viccari is a doctoral student at the LEAD Graduate School & Research Network, which is funded by the Ministry of Science, Research and the Arts of the state of Baden-Württemberg within the framework of sustainability funding for the projects of the Excellence Initiative II. This work was funded by a Fellowship from the Elite Program for post docs of the Baden-Württemberg Foundation and a Jacobs Foundation Research Fellowship (awarded to Lisa Bardach). Lisa Bardach is supported by research funding No P-EDU-23-28, which is co-funded by the European Union (the project "Breakthrough in Educational Research" No 10-044-P-0001) under the 1st April 2025 Agreement with the Research Council of Lithuania (RCOL, LMTLT) and the 14th April 2025 Joint Activity Agreement with Kaunas University of Technology.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.tine.2026.100290.

References

[1] R.E. Mayer, Multimedia learning. Psychology of Learning and Motivation, Academic Press, 2002, pp. 85–139, https://doi.org/10.1016/S0079-7421(02)80005-6.

[2] F. Sella, P. Tressoldi, D. Lucangeli, M. Zorzi, Training numerical skills with the adaptive videogame "the Number Race": a randomized controlled trial on preschoolers, Trends. Neurosci. Educ. 5 (2016) 20–29, https://doi.org/10.1016/j.tine.2016.02.002.

[3] J.L. Plass, B.D. Homer, C.K. Kinzer, Foundations of game-based learning, Educ. Psychol. 50 (2015) 258–283, https://doi.org/10.1080/00461520.2015.1122533.

[4] P. Wouters, H. van Oostendorp, Overview of instructional techniques to facilitate learning and motivation of serious games, in: P. Wouters, H. van Oostendorp (Eds.), Instructional Techniques to Facilitate Learning and Motivation of Serious Games, Springer International Publishing, Cham, 2017, pp. 1–16, https://doi.org/10.1007/978-3-319-39298-1_1.
[5] R.E. Mayer, R. Moreno, Nine ways to reduce cognitive load in multimedia learning, Educ. Psychol. 38 (2003) 43–52, https://doi.org/10.1207/S15326985EP3801_6.
[6] R. Moreno, R.E. Mayer, Cognitive principles of multimedia learning: the role of modality and contiguity, J. Educ. Psychol. 91 (1999) 358–368, https://doi.org/10.1037/0022-0663.91.2.358.
[7] P. Wouters, C. van Nimwegen, H. van Oostendorp, E.D. van der Spek, A meta-analysis of the cognitive and motivational effects of serious games, J. Educ. Psychol. 105 (2013) 249–265, https://doi.org/10.1037/a0031311.
[8] R. Pekrun, The control-value theory of achievement emotions: assumptions, corollaries, and implications for educational research and practice, Educ. Psychol. Rev. 18 (2006) 315–341, https://doi.org/10.1007/s10648-006-9029-9.
[9] R. Pekrun, R.P. Perry, Control-value theory of achievement emotions. International Handbook of Emotions in Education, Routledge/Taylor & Francis Group, New York, NY, US, 2014, pp. 120–141.
[10] S. Hidi, K.A. Renninger, The four-phase model of interest development, Educ. Psychol. 41 (2006) 111–127, https://doi.org/10.1207/s15326985ep4102_4.
[11] S.D. Lipscomb, S.M. Zehnder, Immersion in the virtual environment: the effect of a musical score on the video gaming experience, J. Physiol. Anthr. Appl. Hum. Sci. 23 (2004) 337–343, https://doi.org/10.2114/jpa.23.337.
[12] L.E. Nacke, M.N. Grimshaw, C.A. Lindley, More than a feeling: measurement of sonic user experience and psychophysiology in a first-person shooter game, Interact. Comput. 22 (2010) 336–343, https://doi.org/10.1016/j.intcom.2010.04.005.
[13] J. Byun, C.S. Loh, Audial engagement: effects of game sound on learner engagement in digital game-based learning environments, Comput. Hum. Behav. 46 (2015) 129–138, https://doi.org/10.1016/j.chb.2014.12.052.
[14] D. Kao, R. Ratan, C. Mousas, A.J. Magana, The effects of a self-similar avatar voice in educational games, in: Proc. ACM Hum.Comput. Interact. 5, 2021, pp. 1–28, https://doi.org/10.1145/3474665.
[15] E. Rovithis, A. Floros, A. Mniestris, N. Grigoriou, Audio games as educational tools: design principles and examples, in: 2014 IEEE Games Media Entertainment, IEEE, Toronto, ON, 2014, pp. 1–8, https://doi.org/10.1109/GEM.2014.7048083.
[16] J. McCall, Gaming the Past: Using Video Games to Teach Secondary History, 2nd ed., Routledge, New York, 2023.
[17] S. Domsch, Hearing storyworlds: how video games use sound to convey narrative, in: J. Mildorf, T. Kinzel (Eds.), Audionarratology, De Gruyter, 2016, pp. 185–198, https://doi.org/10.1515/9783110472752-012.
[18] J. McCall, Teaching history with digital historical games: an introduction to the field and best practices, Simul. Gaming 47 (2016) 517–542, https://doi.org/10.1177/1046878116646693.
[19] A. Viccari, R. Göllner, J.U. Hahn, L. Bardach, Limes (the Roman Frontier): developing a video game for history learning, in: ECGBL 18, 2024, pp. 839–847, https://doi.org/10.34190/ecgbl.18.1.2833.
[20] F. Faul, E. Erdfelder, A. Buchner, A.G. Lang, Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses, Behav. Res. Methods 41 (2009) 1149–1160, https://doi.org/10.3758/BRM.41.4.1149.
[21] A. Viccari, A. Kienitz, A. Eitel, J.U. Hahn, R. Göllner, M. Lindner, L. Bardach, Exploring roman history through video game and text: effects of format and instructional framing, (2025). https://doi.org/10.31234/osf.io/fa28t_v1.
[22] A. Schumacher, Y. Kammerer, C. Scharinger, S. Gottschling, N. Hübner, M. Tibus, E. Kasneci, T. Appel, P. Gerjets, L. Bardach, How do intellectually curious and interested people learn and attain knowledge? A focus on behavioral traces of information seeking, Eur. J. Pers. 40 (2026) 87–110, https://doi.org/10.1177/08902070241309124.
[23] L. Linnenbrink-Garcia, A.M. Durik, A.M. Conley, K.E. Barron, J.M. Tauer, S. A. Karabenick, J.M. Harackiewicz, Measuring situational interest in academic domains, Educ. Psychol. Meas. 70 (2010) 647–671, https://doi.org/10.1177/0013164409355699.
[24] M.C. Krebs, K. Braschoß, A. Eitel, Does watching an explainer video help learning with subsequent text? – Only when prompt-questions are provided, Learn. Instr. 94 (2024) 101988, https://doi.org/10.1016/j.learninstruc.2024.101988.
[25] M. Bieleke, K. Gogol, T. Goetz, L. Daniels, R. Pekrun, The AEQ-S: a short version of the achievement emotions Questionnaire, Contemp. Educ. Psychol. 65 (2021) 101940, https://doi.org/10.1016/j.cedpsych.2020.101940.
[26] R.E. Mayer, R. Moreno, Animation as an aid to multimedia learning, Educ. Psychol. Rev. 14 (2002) 87–99, https://doi.org/10.1023/A:1013184611077.
[27] T. Ehrhart, T.N. Höffler, S. Grund, M.A. Lindner, Static versus dynamic representational and decorative pictures in mathematical word problems: less might be more, J. Educ. Psychol. 116 (2024) 532–549, https://doi.org/10.1037/edu0000821.
[28] T. Ehrhart, M.A. Lindner, Computer-based multimedia testing: effects of static and animated representational pictures and text modality, Contemp. Educ. Psychol. 73 (2023) 102151, https://doi.org/10.1016/j.cedpsych.2023.102151.
[29] L. Bardach, K. Murayama, The role of rewards in motivation—beyond dichotomies, Learn. Instr. 96 (2025) 102056, https://doi.org/10.1016/j.learninstruc.2024.102056.
[30] A. Eitel, M.C. Krebs, C. Schöne, Introducing the FoRe-squares model: focus support and reward as key predictors of effective technology-augmented instruction, Educ. Psychol. Rev. 37 (2025) 67, https://doi.org/10.1007/s10648-025-10045-z.
[31] T. Sanders, P. Cairns, Time perception, Immersion and Music in Videogames, BCS Learning & Development, 2010, https://doi.org/10.14236/ewic/HCI2010.21.
[32] S. Erhel, E. Jamet, Digital game-based learning: impact of instructions and feedback on motivation and learning effectiveness, Comput. Educ. 67 (2013) 156–167, https://doi.org/10.1016/j.compedu.2013.02.019.
[33] K. Murayama, L. FitzGibbon, M. Sakaki, Process Account of curiosity and interest: a reward-learning perspective, Educ. Psychol. Rev. 31 (2019) 875–895, https://doi.org/10.1007/s10648-019-09499-9.
[34] S. Erhel, E. Jamet, Improving instructions in educational computer games: exploring the relations between goal specificity, flow experience and learning outcomes, Comput Hum. Behav. 91 (2019) 106–114, https://doi.org/10.1016/j.chb.2018.09.020.